

Poor, or Just Feeling Poor?

On Using Subjective Data in Measuring Poverty

Martin Ravallion

The World Bank
Development Research Group
Director's office
February 2012



Abstract

The challenges faced in calibrating poverty and welfare measures to objective data have long been recognized. Until recently, most economists have resisted a seemingly obvious solution, namely to ask people themselves: “Do you feel poor?” The paper studies the case for and against this approach. It is argued that, while one would not want to use self-assessments as welfare metrics in their own right, there is scope for using such data to help calibrate multidimensional measures. Indeed, the idea

of a “social subjective poverty line” (below which people tend to think they are poor, but above which they do not) is arguably the most conceptually appealing way of defining poverty. However, the paper points to a number of concerns that have received insufficient attention, including the choice of covariates, survey design issues, measurement errors, frame-of-reference effects, and latent heterogeneity in personality traits and personal tradeoffs. Directions for future research are identified.

This paper is a product of the Director’s office, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://econ.worldbank.org>. The author may be contacted at mravallion@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Poor, or Just Feeling Poor?

On Using Subjective Data in Measuring Poverty

Martin Ravallion¹

*Development Research Group, World Bank
1818 H Street NW, Washington DC, 20433, USA*

Key words: Subjective welfare, well-being, multidimensional poverty, scales, vignettes, latent heterogeneity

JEL classifications: C81, D60, I32

World Bank Sector: POV

¹ For comments on the paper the author is grateful to Kathleen Beegle and Nobuo Yoshida. These are the views of the author alone and should not be attributed to the World Bank or any affiliated organization.

1. Introduction

Development monitoring and policy evaluation efforts have traditionally put a high weight on progress against poverty, defined in terms of household consumption or income as measured in surveys.² It is well recognized that household consumption or income aggregates need to be normalized for cost-of-living differences, including differences in household size, although there are differing views on how the normalization should be done. It has also come to be recognized that neither income nor consumption (even with seemingly appropriate normalizations) can be considered a sufficient statistic for welfare. Standard survey methods allow a fairly complete accounting of the market goods consumed (including own-farm products). However, there are important non-market goods that are typically excluded, such as access to public services for health care or schooling. And household aggregate income or consumption reveals nothing directly about distribution within the household.

There have been attempts to define development indices that allow for these “non-income” factors. An early example was Morris’s (1979) “Physical Quality of Life Index”, which is an average of literacy, infant mortality and life expectancy. A now famous example is the “Human Development Index” (HDI), which aggregates life expectancy and education, along with income.³ A recent example is the “Multidimensional Poverty Index” (MPI) of Alkire and Santos (2010). This combines 10 components, two for health (malnutrition, and child mortality), two for education (years of schooling and school enrollment), and six for “living standards” (including both access to services and proxies for household wealth).⁴

All such composite indices confront the problem of setting the trade-offs (relative weights) across their dimensions. This is essentially *ad hoc* in all these measures.⁵ And there are objections that can be raised. For example, Ravallion (2012a) points out that the HDI puts a

² For example, since the mid-1990s the World Bank—the largest international development agency—has deemed poverty reduction to be its “overarching goal.” The first of the United Nation’s Millennium Development Goals is to halve the 1990s income or consumption poverty rate for the developing world by 2015.

³ The index has been published since 1990 in the UNDP’s *Human Development Reports*; see, for example, UNDP (2010).

⁴ For further discussion of the various approaches to measuring multidimensional poverty see Tsui (2002), Bourguignon and Chakravarty (2003), Alkire and Foster (2007), Wagle (2008) and Ravallion (2011).

⁵ Furthermore, the weights are rarely explicit in the space of the postulated core dimensions of welfare. Instead, the weights are only made explicit in a derived space of pre-defined functions of those dimensions. Thus indices such as the HDI and MPI lack transparency about their assumed trade-offs. For further discussion of this point see Ravallion (2011).

remarkably low monetary value on an extra year of life in very poor countries, and a very high value in rich ones.⁶ The HDI's weight on schooling in poor countries is also far higher than its weight on longevity, for no obvious reason. Recognizing that welfare is "multi-dimensional," and that income is an incomplete metric, does not mean we can credibly collapse the multiple dimensions into a single, uni-dimensional, space. The essential problem is that we do not know the prices for valuation. (Indeed, if we knew them we would presumably have already formed a measure of full income.) Prevailing practice is essentially to make *ad hoc* assumptions about the welfare function (and hence the weights).

The main aim of this paper is to review the issues raised by using subjective assessments of welfare (or "well-being"—I use the terms interchangeably) in both deciding who is "poor" and calibrating multidimensional welfare measures. This is not a survey of the literature. Rather the aim is to assess the case for and against the use of subjective data for this purpose and to identify key knowledge gaps.

Scientific research on subjective well-being emerged in the 1960s in psychology (where the emphasis was on measuring and explaining "happiness" as a state of mind) and economics (where the emphasis was on calibrating welfare functions, including setting equivalence scales).⁷ The core data have long been respondents' self-assessments in sample surveys. Each respondent is asked to rate her "economic welfare," or a broader concept such as "satisfaction with life" or "happiness," on an ordinal scale (sometimes called a "Cantril ladder," after Cantril, 1965). Or the respondent provides money metrics of points on qualitative welfare scales, such as the minimum income needed to "make ends meet" (Hagenaars and Van Praag, 1985; Kapteyn et al., 1988) or for multiple points on a ladder (Van Praag, 1968).

However, it would be fair to say that the use of such data in poverty and welfare measurement has stayed at the fringes of practice. In the eyes of some observers, such as Bertrand and Mullainathan (2001), the long-standing scepticism amongst economists about

⁶ The implicit value of an extra year of life in the new HDI introduced by UNDP (2010) varies from \$0.50 per person year in the poorest country to about \$9,000 per person year in the richest (Ravallion, 2012a).

⁷ An early contribution to the quantification of "happiness" in surveys was Cantril's (1965) idea of the "ladder of life." Early applications of subjective welfare concepts in economics included Van Praag (1968), Kapteyn and Van Praag (1976) and Easterlin (1974).

subjective questions is well founded, notably when these data are used as dependent variables. And poverty measurement is a field in which they are dependent variables.

This paper offers a more positive, but still qualified, view on the scope for using subjective data in calibrating poverty measures. There are trade-offs. Subjective data come with problems and the paper tries to identify the main concerns. But it should be acknowledged that these data can provide welfare-relevant information that is simply not available in standard objective data—information that addresses some classic problems of applied welfare economics. After reviewing the economic foundations of the objective approach to poverty measurement in the next section, the bulk of the paper examines the subjective approach. Section 3 sets out some basic theory and compares the two approaches, while section 4 provides examples. Section 5 discusses the problems faced in implementing the subjective approach and some potential fixes.

2. Poverty measurement and non-income dimensions of welfare

There are various “objective” approaches to measuring poverty, including setting poverty lines (Ravallion, 2012b). The approach with firmest roots in economics allows in principle for multiple “non-income” dimensions of welfare, although these enter in a specific way, namely as shifters of the “basic needs” demand vector needed to attain the poverty-level of welfare. This section reviews this approach, as a prelude to discussing the potential role of subjective data.

The core idea is a utility function for individual i of the form $u(q_i, x_i)$, where q_i is a vector of the quantities of commodities consumed, and x_i is a vector of welfare-relevant “non-income” characteristics, including demographic characteristics of the household. The utility maximizing consumption vector is denoted $q(p_i, y_i, x_i)$ at price vector p_i and total expenditure on consumption y_i . The implied indirect utility function is $v(p_i, y_i, x_i)$, giving the maximum attainable welfare at prevailing prices and characteristics. On inverting we obtain the expenditure function $e(p_i, x_i, u)$ giving the minimum cost of utility u for person i when facing prices p_i .

In using this standard model for poverty measurement, poverty is thought of as being absolute in the space of welfare. The minimum utility that is deemed necessary to escape poverty is denoted u_z . The welfare-consistent poverty lines are then given by:

$$z_i^u = e(p_i, x_i, u_z) \quad (i = 1, \dots, N) \quad (1)$$

On exploiting the standard properties of the expenditure function, this can be re-written as:

$$z_i^u = p_i q^c(p_i, x_i, u_z) \quad (2)$$

where $q^c(p_i, x_i, u_z)$ is the vector of utility-compensated demands. Thus we see that the welfare-consistent poverty line is the cost of a bundle of “basic consumption needs,” given by the vector of utility-compensated demands at the reference level of utility defining who is poor in the welfare space. The measure of aggregate poverty can then be defined on the vector of welfare ratios, $\{y_i / z_i^u, i=1, \dots, N\}$.⁸ The “poverty rate” or “headcount index of poverty” is the proportion of the population with $y_i / z_i^u \leq 1$.⁹

Three remarks are in order. First, notice that, even though individuals are ranked by their consumption expenditure, the implied poverty rate is consistent with that implied by the multi-dimensional welfare function, $u(q_i, x_i)$. This comes from the fact that a person is “welfare poor,” meaning that $u_i \leq u_z$, if and only if she is consumption poor, defined as $y_i \leq z_i^u$. (This follows immediately from the fact that the expenditure function is the inverse of the indirect utility function.) The fact that one uses an income poverty measure does not mean that one thinks of welfare as only depending on income. However, for income-based poverty measurement to be welfare consistent it is imperative that the poverty lines are adjusted for all relevant non-income dimensions of welfare. Adjustment for prices is common, but the only other adjustments typically done are for (almost invariably) household size and (sometimes) demographic composition. It is in the practice of this approach that concerns about omitted “non-income” dimensions of welfare become salient.

Second, note that this framework allows for a measure of absolute income poverty. One can readily allow for a relative poverty line by supposing that the vector x includes mean income

⁸ The term “welfare ratios” is due to Blackorby and Donaldson (1987) who discuss the advantages of this welfare metric over money-metric utility or “equivalent income functions.”

⁹ A well-known problem with this measure is that it ignores distribution below the line; for example, if the poorest person becomes worse off or an inequality-increasing transfer is made amongst the poor then the headcount index is unchanged. There is a large literature on alternative measures; for a survey see Zheng (1997).

of some reference group. If people derive utility from relative position then an absolute poverty line in the space of welfare will translate into a relative line in the space of commodities (Ravallion, 2008b).

Third, the basic model above can be extended to encompass Sen's (1985) argument that welfare should be thought of in terms of a person's capabilities—the functionings (“beings and doings”) that a person is able to achieve. By this view, poverty means not having an income sufficient to support specific normative functionings.¹⁰ A more general definition of the poverty line is proposed in Ravallion (1998, 2012b) whereby a person's functionings are determined by the goods she consumes and her characteristics. (Functionings may also vary with the mean consumption of an appropriate reference group.) One can further postulate that utility is derived directly from functionings and interpret $u(q_i, x_i)$ as a derived utility function, after solving out the functionings. Again, individual choices of q_i maximize utility. Thus one can also postulate solutions for functionings, showing how they depend on income, prices and characteristics. “Functioning consistency” for a set of poverty lines requires that certain normative functionings are reached at the poverty line.¹¹

The real challenges are in putting these ideas into practice. This requires that we address two classic problems in applied welfare economics. The first can be called the referencing problem: what is the reference level of utility (or other functionings) that anchors the poverty line? (This is an instance of a general problem of reference-dependence, including in welfare measurement.) It is tempting to say this choice is arbitrary, and to hope that it is innocuous. But the choice of the reference in poverty measurement is far from arbitrary, since (in general) it affects the resulting poverty measure. This speaks to the importance of testing the sensitivity of poverty comparisons to the choice of reference, as it determines the level of the poverty line.¹²

The second problem is the identification problem. Even if we agree on the poverty line in the welfare space, we also need to know the expenditure function in (1). Standard practice is to calibrate its parameters from data on consumer demands. The problem is that individuals vary in

¹⁰ For example, following Atkinson and Bourguignon (2001), one can think of the normative functionings as comprising “survival needs” and minimum “social inclusion costs” for participating in social and economic activity.

¹¹ A more complete discussion can be found in Ravallion (2012b).

¹² This idea is made explicit and formalized in the use of stochastic dominance criteria for ordinal poverty comparisons (Atkinson, 1987; Foster and Shorrocks, 1988).

characteristics, such as their household size and demographic composition, which can influence welfare in ways that may not be evident in consumer demand behavior. More precisely, if demands $q(p_i, y_i, x_i)$ can be integrated back to a utility function $u(q_i, x_i)$ then the same demands also maximize the broader welfare function $w[u(q_i, x_i), x_i]$. Thus there is no unique utility function that can be inferred from market demands alone.¹³ We need stronger identifying assumptions, or more data. The rest of this paper will consider the scope for using data on self-assessed welfare for addressing these problems.

3. Anchoring poverty measures to subjective welfare

There are essentially two approaches to collecting subjective data on welfare. The first uses qualitative categories in the welfare space. An example is the “economic ladder question” (ELQ): *“Imagine six steps, where on the bottom, the first step, stand the poorest people, and on the highest step, the sixth, stand the rich (show a picture of the steps). On which step are you today?”* A popular indicator of overall welfare has been the “satisfaction with life” (SWL) question: *“Overall, how satisfied (content, happy) are you with your life? Are you (1) very unsatisfied; (2) unsatisfied; (3) neither unsatisfied nor satisfied; (4) satisfied; (5) very satisfied?”* The ELQ is arguably better suited to poverty measurement as it is tied to a concept of economic welfare, while ideas such as SWL or “happiness” are broader and more nebulous in scope.

The second approach asks for a money-metric of subjective welfare. An example is the “minimum income question” (MIQ): *“What income level do you personally consider to be absolutely minimal? That is to say that with less you could not make ends meet.”* This can be thought of as a special case of Van Praag’s (1968) “income evaluation question,” which asks what income is considered “very bad,” “bad,” “not good,” “not bad,” “good,” “very good.”

There are three main ways such data have been used to inform poverty measurement. The first is as a means of testing objective poverty lines, by regressing self-rated welfare on income normalized by the poverty line *plus* the variables that went into the construction of the poverty line, which should be jointly insignificant if those lines accord with subjective welfare.¹⁴ The second is to calibrate a composite welfare index using the subjective welfare regression

¹³ This point has long been recognized; see Pollak and Wales (1979), Pollak (1991) and Browning (1992).

¹⁴ This approach is outlined in Ravallion and Lokshin (2002) and illustrated using Russia’s poverty lines.

coefficients as the weights.¹⁵ The third way is to derive a poverty line in the income space, defined as the income level at which some critical level of subjective welfare is reached in expectation. Notice that, as long as higher income is valued positively in subjective welfare, the latter two methods are equivalent in terms of their implied poverty rate (proportion of the population living in poverty) for any given poverty line in the space of subjective welfare. As noted in the previous section, the fact that one uses income as the ranking variable does not mean that one is ignoring non-income dimensions of welfare given that these are allowed to alter the poverty line consistently with their welfare weight.¹⁶

It is well-recognized that there are differences in self-assessments that do not pertain to economic welfare. One might define as poor all those who put themselves below some fixed rung of the ELQ or all those with income less than their answer for the MIQ. However, one must allow for heterogeneity, such that people at the same standard of living give different answers to the subjective welfare question, or that the same person may give a different answer depending on when they are interviewed. These variations could stem from idiosyncratic “mood” effects, personality traits, or simply errors. It is not one’s stated perception of own-welfare in an interview that is taken to be the relevant metric, as this is bound to contain many factors that one would not deem relevant. Rather, the subjective questions are used to calibrate an inter-personally comparable welfare function based on observed covariates deemed to be relevant on *a priori* grounds. It is the role of these data sources in calibration that is the present focus.

Let us take a closer look at one approach. This recognizes that any definition of poverty must be broadly accepted in the specific context it is being used. It makes little sense to use a poor country’s poverty line in a rich country; indeed, it is unlikely that anyone in the US or Western Europe lives below the poverty lines found in the poorest countries, at around \$1 a day. And if one used the US poverty line in India (say) one would find that all but 5% of the population is poor (Ravallion, 2010); it is extremely unlikely that most Indians would agree, and nor does a poverty rate of 95% help much in thinking about how best to fight poverty in India. The idea of a “social subjective poverty line” (SSPL) is a natural way of defining poverty in a

¹⁵ Examples are found in Ferrer-i-Carbonell and Van Praag (2001) and Kingdon and Knight (2006).

¹⁶ This point is a common source of confusion in the more applied literature and policy-related discussions on multidimensional poverty measurement where “income poverty” is often counterpointed to “multidimensional poverty.” Plainly the two approaches are equivalent when calibrated to the same welfare function.

specific context. The SSPL can be defined as the income below which people tend to think they are poor in the specific setting and above which they tend to think they are not poor.

The standard representation of this approach in the literature is essentially a statistical model of the following form. Let y_i^{\min} denote the answer given by person i to the MIQ and write:

$$y_i^{\min} = E(y_i^{\min} | p_i, y_i, x_i) + \varepsilon_i \quad (3)$$

where $E(y_i^{\min} | p_i, y_i, x_i)$ is the conditional mean—the regression function—given actual income y , which lies in the interval $[y_0, y_1]$, prices and other relevant characteristics, and ε_i is a zero-mean error term. It is assumed that $E(y_i^{\min} | p_i, y_i, x_i)$ is strictly increasing in y . It can also be assumed that $E(y_i^{\min} | p_i, y_0, x_i) > y_0$ and $E(y_i^{\min} | p_i, y_1, x_i) < y_1$, in which case it is plain that there exists a unique fixed point such that:

$$E[y_i^{\min} | p_i, z(p_i, x_i), x_i] = z(p_i, x_i) \quad (4)$$

The social subjective poverty line (SSPL) is then the (unique) solution of (4) for $z(p_i, x_i)$. If higher p or x is associated with higher y^{\min} then a higher SSPL will be required.

The conceptual links between the objective, utility-based, approach (section 2) and the two approaches to subjective welfare—the direct approach of asking the respondent’s position on a welfare ladder and the approach of asking for a monetary valuation—have never been clear in the literature, and the three approaches have largely existed independently. It is of interest to ask under what conditions they will be consistent in terms of whom they identify as “poor.” Clearly this requires that all three are anchored to the same continuous welfare function. It is evident that the direct approach will be consistent with the objective utility-based approach as long as the ordinal qualitative questions asked of respondents stem from the same welfare index. This is a judgment call, and some observers may not accept that (say) SWL is driven by “utility.”

Establishing consistency for the SSPL approach requires that we introduce some economic structure into the statistical model of the SSPL, as above. We can assume that the MIQ attempts to find y_i^{\min} such that:

$$v(p_i, y_i^{\min}, x_i) = u_i^r \quad (5)$$

where u_i^r is the level of utility that person i thinks is implied by “making ends meet.” This is assumed in turn to be a strictly increasing function of actual utility:

$$u_i^r = \phi[v(p_i, y_i, x_i)] \quad (6)$$

Equation 3 can now be interpreted as the solution of (5) and (6) for y_i^{\min} . The SSPL in the income space is then found as the solution to:

$$v(p_i, z_i, x_i) = \phi[v(p_i, z_i, x_i)] \quad (7)$$

But now we see that there is also a SSPL in welfare space, u_z , at which point:

$$u_z = \phi(u_z) \quad (8)$$

(The solution is unique if $\phi(u_0) > u_0$ and $\phi(u_1) < u_1$, where utility is bounded below by u_0 and above by u_1 .) Comparing (7) and (8) we see that $u_z = v(p_i, z_i, x_i)$. The SSPL—the point at which $y_i^{\min} = y_i$ —also identifies those who just reach the implicit welfare poverty line u_z and the set of people who are welfare poor ($u_i < u_z$) is the same as the set of income poor ($y_i < z_i$). Thus, under these conditions, the SSPL approach can be said to be consistent with the other two approaches, in that they identify the same people as poor and deliver the same poverty rate.

4. Methods and examples

The most common method of empirical investigation in this literature has been a regression of the survey responses on subjective welfare on individual and household characteristics, including age, gender, income, education, employment status and household demographics.¹⁷ Typically the researcher has data on subjective welfare (SW) across N survey

¹⁷ Examples include van Praag (1968), van de Stadt et al. (1985), Clark and Oswald (1994, 1996), Kapteyn et al. (1998), Easterlin (1995), Oswald (1997), Winkelmann and Winkelmann (1998), Pradhan and Ravallion (2000), McBride (2001), Ravallion and Lokshin (2001, 2002, 2010), Graham and Pettinato (2002), Senik (2004), Luttmer (2005), Ferrer-i-Carbonell (2005), Herrera et al. (2006), Bishop and Luo (2006), Kingdon and Knight (2006, 2007), Fafchamps and Shilpi (2009), Knight and Gunatilaka (2010) and Castilla (2010).

respondents and a set of covariates (X , including some suitable non-linear function of income) and postulates a regression model of the form:

$$SW_i = \alpha + \beta X_i + \varepsilon_i \quad (i=1,...,N) \quad (9)$$

where the error term, ε , is uncorrelated with X . If using ordinal data from a ladder question then an ordered response estimator is typically used, such as the ordered probit, in which case (9) is the underlying continuous variable generating the ordinal responses relative to thresholds that are taken to be fixed. If instead one is using the MIQ, then y_i^{\min} is the dependent variable.

This section provides an overview of some of the lessons to date in estimating such a model. The next section will discuss a number of issues in estimating such a regression.

Recall that one of the conditions for existence of a SSPL is that subjective welfare is an increasing function of income. This is invariably indicated in cross-sectional studies; indeed, I have not seen an exception in the literature though some studies find stronger income effects than others.¹⁸ Stevenson and Wolfers (2008) argue that the income effect on subjective well-being is robust across countries, within countries, and over time.

There have been some estimates of SSPLs.¹⁹ Interestingly, the estimates to date suggest that the overall poverty rate based on the SSPL is roughly similar to that implied by objective poverty lines.²⁰ It may well be that the choice of parameters in the “objective” absolute lines already approximated the expected SSPL in the specific context.

However, the structure of the poverty profile has turned out to be different in some respects. While objective poverty lines often imply that larger households are poorer, this is not typically the case in cross-sectional studies using the subjective approach, which tends to suggest greater economies of scale in consumption than normally assumed, at least in developing countries. For example, in using the ELQ to test the welfare consistency of prevailing objective

¹⁸ Clark et al. (2008) review the literature on the income effect on subjective welfare.

¹⁹ I focus on application to developing countries. The applications to date include Pradhan and Ravallion (2000) using data for Jamaica and Nepal, Ferrer-i-Carbonell and Van Praag (2001), for Russia, Taddesse and Shimeles (2005) for Ethiopia, Gustafsson (2004) for urban China, Lokshin et al. (2006) for Madagascar, Bishop and Luo (2006) for urban China and Carletto and Zezza (2006) for Albania.

²⁰ An exception to this finding is reported for the US by de Vos and Garner (1991), where the SSPL is well above the prevailing (absolute) line, though the US line has not been updated in real terms since the 1960s; a more current absolute line for the US would probably be closer to the SSPL.

poverty lines for Russia, Ravallion and Lokshin (2002) find some striking differences in the properties of the equivalence scale. The objective poverty lines had an elasticity of 0.8 to household size, while the subjective indicator called instead for an elasticity half this size.²¹

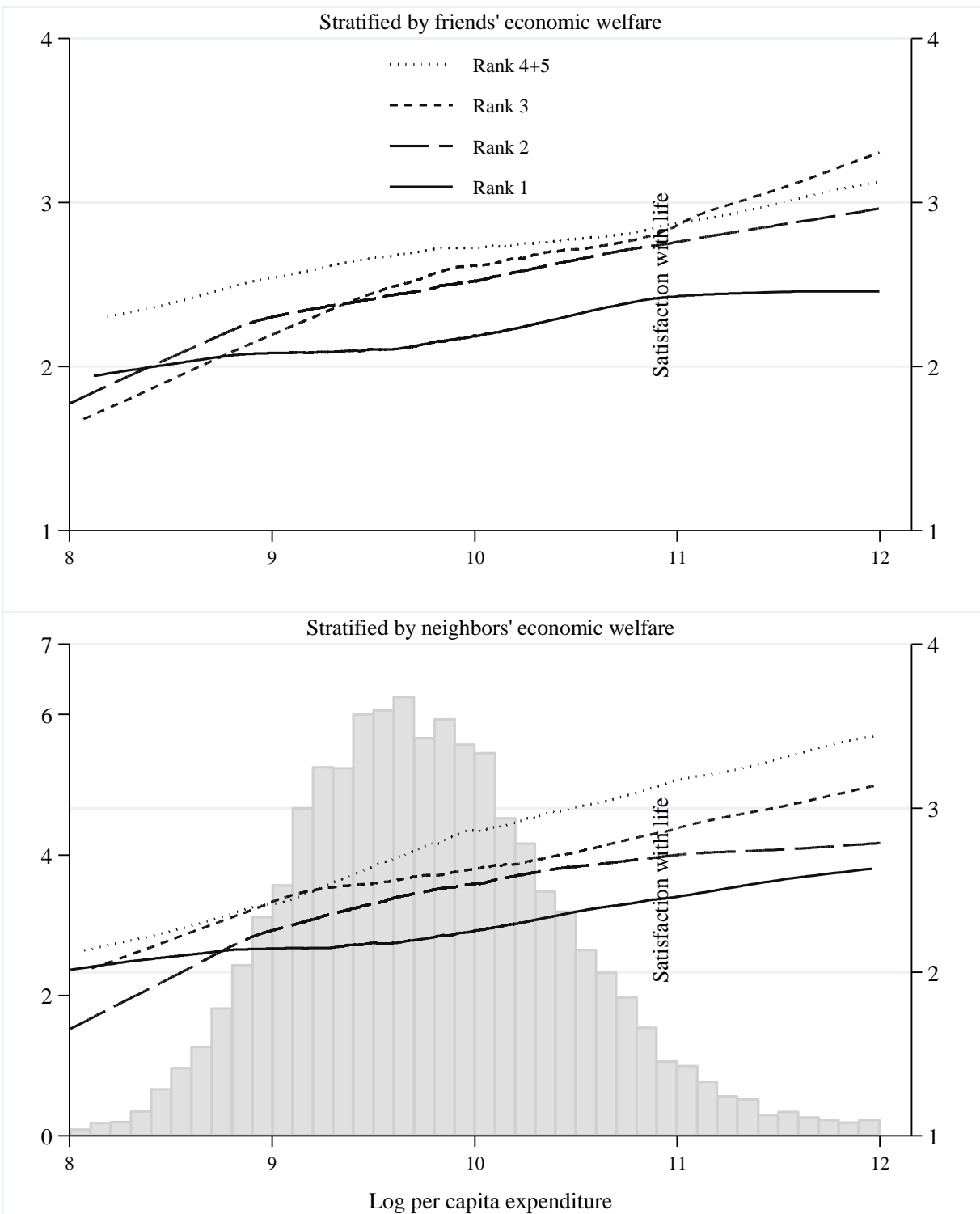
Subjective data have thrown new light on the long-standing debate on whether poverty is “absolute” or “relative.” The position one takes on this issue is crucial to whether one believes that economic growth reduces poverty. If the poverty line has an elasticity of zero to the poverty line then growth tends to reduce poverty, but this ceases to be true at an elasticity of one (Ravallion, 2003). However, one finds little support for the idea of a relative poverty line set at a constant proportion of the current mean income. Poverty lines calibrated to subjective welfare tend to rise with mean income but with an elasticity less than unity, suggesting that they are more like the “weakly relative poverty lines” defined by Ravallion and Chen (2011). Hagenaars and Van Praag (1985) estimated an elasticity of 0.51 for eight European countries. For the US, Kilpatrick (1973) estimated an elasticity of about 0.6 for subjective poverty lines and De Vos and Garner (1991) found an own-income elasticity of the US subjective poverty line of 0.43.

A number of papers have reported evidence of effects on subjective welfare that can be interpreted as indicative of “relative deprivation,” meaning that self-assessed well-being tends to fall as social comparators become better off, at given “own income.”²² Luttmer (2005) reports regressions for subjective welfare in the US that imply a particularly strong relativism, whereby own income does not matter to subjective well-being independently of income relative to the mean in the area of residence. (The next section points to some reasons to question this finding.) The bulk of the evidence has been for relatively rich countries. The work that has been done for developing countries has been less supportive. Senik (2004), Kingdon and Knight (2007) and Ravallion and Lokshin (2010) tested for relative deprivation effects in self-reported happiness in Russia, South Africa and Malawi respectively and found rather little support for the idea and even evidence of positive external effects of higher “neighbors’ income,” rather than the negative effect predicted by the theory of relative deprivation.

²¹ Similarly, see Pradhan and Ravallion (2000), using data for Jamaica and Nepal, Bishop and Luo (2006), using data for urban China, and Rojas (2007) using data for Mexico. For a more general discussion of economies of scale in consumption in developing countries see Lanjouw and Ravallion (1995).

²² See Oswald (1997), Frank (1997), Frey and Stutzer (2002), Clark, Frijters, and Shields (2008) and Clark and Senik (2010). Reviewing the evidence, Frey and Stutzer assert that “There is little doubt that people compare themselves to other people and do not use absolute judgments” (p.412). This would seem to be over-stated.

Figure 1: Nonparametric regressions of satisfaction with life against log consumption per person in Malawi, 2004



Source: Ravallion and Lokshin (2010).

As an illustration, Figure 1 shows the relationship found by Ravallion and Lokshin (2010) in survey data for one of the world's poorest countries, Malawi. The Figure gives the non-parametric regression for responses to the “satisfaction with life” question against household consumption per person. The Figure also shows how the relationship varies with the respondent's assessment of the economic welfare of their friends and neighbors. We see that SWL increases with consumption, but higher economic welfare of friends or neighbors tends to come with higher satisfaction with life at given own consumption. So this does not support the theory of relative deprivation (as this would imply lower welfare with richer friends or neighbors, at given own consumption). The parametric models presented in Ravallion and Lokshin (2010) did, however, indicate that concerns about relative deprivation emerged amongst the richer Malawians. Ravallion (2008b) calculated the implied (iso-welfare) poverty lines for Malawi, and showed that they exhibit a small weakly relativist gradient, whereby higher own-economic welfare is needed to compensate people for having richer social comparators.

These findings are echoed in the responses to survey questions on whether poverty is perceived as absolute or relative reported by Corazzini et al. (2010), which suggest that respondents coming from richer backgrounds see poverty as more relative, though not to the point that absolute poverty is seen as unimportant.

Some of the literature has also thrown light on the role played by expectations. Ravallion and Lokshin (2002) find strong lagged income effects on subjective welfare in Russia, consistent with the view that those on a positive (negative) income trajectory are likely to rate their welfare higher (lower) than others at given current income (and other covariates).²³

Knight and Gunatilaka (2012) argue that “aspirations” for higher income also matter negatively to subjective welfare, and that this mitigates appreciably the direct income effect. Using data for rural China they proxy “aspiration income” by answers to a version of the MIQ and find that the latter has a significant negative coefficient in predicting SWL and “satisfaction with income.” Aspiration income is in turn found to be an increasing function of own income (consistently with past research on the SSPL). They interpret their results as evidence of an “hedonic treadmill” whereby higher actual income need not raise happiness given offsetting

²³ Ravallion and Lokshin (2000) also find that Russians on a positive (negative) income trajectory are also less (more) likely to support redistribution from the rich to the poor.

effects on aspirations.²⁴ However, once one allows for latent heterogeneity in subjective welfare there are other interpretations of the Knight-Gunatilaka finding, as discussed in the next section.

A number of papers have found evidence that unemployment lowers subjective welfare at given income.²⁵ This is not what the standard economic model of work-leisure choice would suggest, since unemployment at given income implies greater leisure, which is assumed to yield utility. (The welfare cost of unemployment is seen to be through the lost income.) There may well be an independent disutility of unemployment that is missing from the standard model, possibly associated with the quantity constraints on choice that involuntary unemployment entails or possibly stemming from social status roles played by employment. However, the welfare effect attributed to unemployment may also stem from latent heterogeneity in personality traits, which is one of the issues to which the next section turns.

5. Problems in using subjective data in measuring poverty

Subjective data can offer a direct lens on welfare that is not available in standard objective data. But are subjective questions reliable, in the sense that one gets similar answers under similar circumstances? Studies of SWL or “happiness” have found a worryingly high degree of variability. For example, Kristensen and Westergaard-Nielsen (2007) find that 20% of survey respondents in Europe gave a different answer on job satisfaction when asked twice within the same interview. Krueger and Schkade (2008) studied reported SWL for the same people interviewed two weeks apart and found the serial correlation coefficient was only 0.6.

In the present context, it is not the reliability of the subjective welfare scores themselves that is of greatest concern, but the reliability of the regressions for subjective welfare, as used in calibrating social-subjective poverty lines, or measuring welfare more generally. If the likely errors in subjective welfare data are white noise—classical measurement errors—we would not be concerned; the R^2 will be lower but the regression coefficients will still be unbiased under standard assumptions. However, there are good reasons to expect the errors in subjective data to be correlated with other explanatory variables of interest, as argued by Bertrand and

²⁴ The authors note the endogeneity concern and provide instrumental variables (IV) estimates although it is not obvious that there is any theoretically valid IV for this problem.

²⁵ Examples include Clark and Oswald (1994), Theodossiou (1998), Winkelmann and Winkelmann (1998) and Ravallion and Lokshin (2001).

Mullainathan (2001) and Ravallion and Lokshin (2001). That does not mean the regressions are useless. Biases can almost never be ruled out. The real issue is how large the biases are in practice, and whether they can overturn key conclusions, such as whether larger households are poorer or whether economic growth reduces poverty.

This section tries to assess what appear to be the main concerns about the specification and interpretation of regressions for subjective welfare.

Welfare-irrelevant covariates and missing data on welfare-relevant ones

To assert that subjective welfare regressions should over-ride prior normative judgments puts a lot of weight on those regressions. There are concerns about both the variables included and the variables left out.

Not everything that helps predict self-assessed welfare will be considered relevant to measuring poverty (or assessing welfare more broadly). In principle one can partition the X vector in equation (9) into two types of variables, those that are deemed (on *a priori* grounds) to be relevant to assessing welfare, and those that are controls for other factors, with the latter fixed at reference values (common to all individuals) when measuring welfare or poverty.²⁶ Thus the welfare metric becomes a predicted value as if the individual had the reference characteristics.

Personality traits are a case in point. These clearly influence self-assessed well-being. While these differences are not often data, this is changing. Suppose that we control for personality in a regression for (say) SWL. The idea of a poverty line that is set at a higher level for people with certain personalities, at given levels of other relevant factors, is probably not going to gain much acceptance. Personality is unlikely to be given independent weight.

And some more readily observed covariates may well be acting as proxies for personality. For example, a common finding in the literature is that unemployment reduces subjective welfare at a given level of income. If one included unemployment in defining the poverty line then one would conclude that the unemployed should have a higher poverty line than the employed, *ceteris paribus*. Yet other “objective” approaches to setting a poverty line

²⁶ This is proposed by Ravallion (1990) in the context of measuring undernutrition, though the analytics are formally identical.

would not suggest that the unemployed need more income to attain the same level of welfare. And that may be the right call if the trade-off between income and unemployment implied by a subjective welfare regression is being distorted by latent personality traits.

Nor are the data for X available ever likely to embrace all that matters to welfare. This approach is unlikely to provide a sufficient statistic for multidimensional poverty comparisons. (And it is unlikely that any approach would.) To estimate the regression function we ideally have the X 's from the same individuals as the indicator of subjective welfare. There is bound to be some relevant dimension of welfare left out, but hopefully available in some separate survey. In practice, multidimensional poverty measurement is invariably going to be a matter of devising a defensible set of multiple indices rather than one single composite index (Ravallion, 2011).

Survey design and measurement errors

Measurement errors are present in all areas of applied research. Naturally there are errors in survey responses on subjective welfare or its determinants. Three examples follow of why such errors matter to the conclusions drawn about the determinants of subjective welfare.

(i) Incomes or consumptions are measured with error. Sometimes the income measure was built up from a detailed questionnaire, which is likely to be more reliable than other measures found in the literature based on very few questions, or even just one, “what is your income?” The scope for attenuation bias in estimating the income gradient in subjective welfare is plain.

(ii) An attenuation bias in estimates of the income (or other) gradient in subjective welfare can also arise from psychological adaptation to adverse circumstances. Poor people may try to cope with deprivation in ways that attenuate the true income gradient in subjective assessments of their well-being. For example, Lokshin and Ravallion (2008) found that the income gradient in the self assessed health status of Russians was greatly attenuated relative to that in their objective health status.

(iii) Recall that a common method of testing for welfare costs of relative deprivation is by regressing self-assessed welfare on both own income and the mean income of “neighbors”—typically the mean income of survey respondents in a geographic area that includes the

respondent. However, the latter could just be picking up a spatial autocorrelation in income measurement errors, giving spurious social effects.

The measurement of comparison-group welfare also poses a problem. In the near-universal approach in the literature, the researcher must make an *a priori* judgment about the relevant comparison group. Yet this can vary greatly from person to person. It has been argued that "... comparisons are most salient if individuals perceive the reference person or group as in some way similar to themselves" (Kahneman and Varey, 1991, p. 140). But that is based on the individual's own judgment, given idiosyncratic informational and social factors that a researcher would have a very hard time observing in any systematic way. Neither psychological nor economic theories of relative deprivation offer much insight into what constitutes the relevant comparison group. The researcher must make some potentially strong identifying assumptions.

There are also concerns about survey design and implementation effects, including sensitivity to how questions are asked and the context. It is known that precisely how, and in what circumstances, one asks subjective questions matters to the results. In survey data for Britain, Taylor (2006) finds that those interviewed on a Friday report significantly greater job satisfaction and less mental stress than those interviewed mid-week, controlling for other observed factors. Also for Britain, Conti and Pudney (2011) find that seemingly small changes in how questions on satisfaction with work were asked led to large changes in the answers obtained, particularly for women. Moreover, they find that the measurement errors are not ignorable, in that the distortions influence findings on the determinants of job satisfaction. They also find evidence of a "put a good show for the visitor effect" (Conti and Pudney, 2011, p.1093)—whereby there is an upward bias in reported welfare in open oral interviews relative to when the questionnaire is completed in private—and a "not in front of the children effect" (p.1093), whereby there was another upward bias when children were present at the interview. These biases were found to be larger for women. Bertrand and Mullainathan (2001) give other examples of how the ordering and wording of subjective questions can alter responses.

Such problems are not confined to subjective questions. Seemingly objective questions can entail similar biases. My favorite (undocumented) example is from a colleague who was doing field work piloting a questionnaire and found a markedly higher literacy rate for a given population and date when asking "are you literate?" than when using a direct test by showing one

sentence on a card and asking what it said. (In this case the bias was greater for men.) While we do not normally think of literacy or schooling as subjective, such variables are not immune to the same biases.

Some of the long-standing questions on subjective welfare in the literature seem problematic when one considers the scope for measurement error. While the Minimum Income Question offers an elegant solution to the problem of determining the SSPL, it comes with a problem, and it is a problem that is likely to be especially serious in developing-country settings. The problem is that the income concept respondents to the MIQ have in mind need not correspond to that used by the analyst in estimating the SSPL. This will bias the SSPL (Kapteyn et al., 1988; Tummers, 1994). In applying the MIQ in many developing countries, one will also find that “income” is not a well-defined concept, particularly (but not only) in rural areas. It is not at all clear whether one could get sensible answers to the MIQ.

To address this concern, Pradhan and Ravallion (2000) propose a method for estimating the SSPL based on qualitative data on consumption adequacy. Instead of asking respondents what the precise minimum consumption is that they need, Pradhan and Ravallion simply asked whether their current consumptions are adequate. For example, their adequacy question for food takes the form:

Concerning your family's food consumption over the past one month, which of the following is true?

- | | |
|---------------------------|---|
| <i>Less than adequate</i> | 1 |
| <i>Just adequate</i> | 2 |
| <i>More than adequate</i> | 3 |

"Adequate" means no more nor less than what you consider to be the minimum consumption needs of your family.

The question is then repeated for other components of consumption (for both market and non-market goods and services).

This provides a multidimensional alternative to the one-dimensional and difficult MIQ. The method exploits the structure of consumption—essentially providing a subjective welfare function for each component of basic needs. The SSPL can be defined as the level of total spending above which respondents say (on average) that their expenditures are adequate for all their needs or some sub-set. Under certain technical conditions, a unique solution for the SSPL

can then be obtained from the estimated parameters of a set of regressions for consumption adequacy; Pradhan and Ravallion (2000) provide a detailed exposition.

Non-ignorable latent heterogeneity

By “latent heterogeneity” I mean that observationally similar people give different responses to questions on subjective welfare. These differences are “non-ignorable” statistically if they are correlated with the dimensions of welfare for which one wants to derive weights. There are a number of reasons to be concerned about such heterogeneity. It has been argued by psychologists and economists that relatively stable (though not necessarily fixed) personality traits influence behavior in a wide range of domains, including behaviors relevant to economic welfare.²⁷ It appears likely that latent personality traits would be correlated with the observed socioeconomic characteristics that are often postulated as relevant to a multidimensional welfare index, as argued by Ravallion and Lokshin (2001). People with a latent tendency for depression may well be less likely to take care of their physical health. Unemployment may also be more likely for people with certain personality traits, such one’s “work ethic.” These differences can be thought of as either differences in the meaning attached to the scales used in subjective welfare questions or a difference in the latent subjective welfare itself.

To give a sharp example, suppose subjective welfare depends on income and personality, and we are interested in identifying the income poverty line corresponding to some fixed level of subjective welfare. Personality is unobserved (and so relegated to the regression error term), but is known to also affect earnings and (hence) income.²⁸ It is readily verified that if the combined effect of the latent personality traits that raise subjective welfare is to increase (decrease) labor market earnings then we will underestimate (overestimate) the SSPL.

We may want to control for personality in measuring poverty, but we can’t if it is not data. For example, we can agree that people with fewer opportunities for schooling are poorer at given current income. In data we see that schooling is associated with higher subjective welfare at given current income (possibly reflecting higher expected future income for those with more schooling). Thus, welfare consistency appears to demand that we use a higher income poverty

²⁷ See, for example, Goldberg (1993) and Heckman (2011). These traits can change over a life time, but can be treated as roughly constant over periods of a few years or more.

²⁸ Evidence on this can be found in Mueller and Plug (2006). Also see the discussion in Heckman (2011).

line for people with less schooling. However, schooling attainments depend on individual personality as well as opportunities. For example, there is evidence that more conscientious people acquire more schooling.²⁹ At the same time, more conscientious people tend to delay gratification in other aspects of their life, which may well mean that they report lower current subjective welfare at given current income than do others. Then a regression for subjective welfare ignoring the personality differences will tend to lead one to underestimate the required gradient in poverty lines needed to compensate for differences in schooling opportunities.

Even with similar personalities (or statistically ignorable differences) different people may apply different criteria for scaling their welfare. They have different ideas about what it means to be “rich” or “poor,” or what it means to be “satisfied” or not with one’s life. Again latent heterogeneity in factors that are essentially irrelevant to welfare but influence responses to subjective welfare questions can cast doubt on the implied interpersonal comparisons of welfare from subjective data. If such heterogeneity was purely random then it would not invalidate inferences from such regressions (at least for linear models). Thus it has been argued in the literature that, while inter-personal welfare comparisons are invalidated by heterogeneous scales, the regressions are likely to be robust to such heterogeneity.³⁰

However, that claim is of doubtful veracity. As already noted, it is a well-established idea that people assess their welfare relative to some “comparison group” such as neighbors or co-workers. This argument has emphasized relativist welfare comparisons. It has also been argued that reference groups play an important role in expectations formation.³¹ It is a small step from these ideas to the proposition that survey respondents answer questions with reference to their immediate experiences and that this may well be highly localized in some relevant social or geographic dimensions. People will answer subjective questions relative to their personal “frame-of-reference,” which depends on the respondent’s own knowledge and experience, and therefore is likely to vary systematically with the characteristics of that person, including objective measures of economic welfare.

²⁹ Reported in Heckman (2011), using data for Germany.

³⁰ For example, Frey and Stutzer (2002) note the possibility of heterogeneity in the scales used in self-reported welfare questions but claim that this does not invalidate regression models for such data.

³¹ For more recent discussions and evidence see Ravallion and Lokshin (2000) and Senik (2004).

For example, it can be conjectured that people living in poor areas of a developing country tend to have a more limited knowledge of the full range of levels of living found in the society as a whole. Someone living in a poor, remote village who has only infrequently left the village and gone no further than the district town is likely to rate her welfare higher than someone with the same real income who lives in a city and sees far greater affluence around her. Similarly, it can be conjectured that well-off people are often unaware of how poor some people are, and may thus rate their own welfare lower on the scale.

This can be called a frame-of-reference bias (FORB) (Ravallion, 2008b). The potential for FORB raises concerns about the (enumerable) regression models found in past literature. Consider, for example, the many papers that have used subjective welfare regressions to test for reference-group effects, such as whether higher neighbors' income makes one feel worse off through perceptions of relative deprivation. It seems likely that the same reference group also influences the respondents' interpretation of scales used in subjective questions. The reference group acts as both the comparator in assessing relative position and a key element of the information set used by respondents when interpreting the scales.

So a cloud of doubt hangs over subjective welfare regressions, arising from concerns about likely heterogeneity in the interpretation of the scales used to measure subjective welfare. This heterogeneity undoubtedly reduces the power of standard covariates in explaining perceived welfare. More worrying, however, is the possibility that the heterogeneity in scales is leading to biased inferences from studies of subjective welfare, including biases in identifying its income gradient, as well as the effects of other variables such as education, employment status and relative economic position. Bias arises if these variables are correlated with the latent heterogeneity in scales. It can be conjectured that poorer people tend to have more limited horizons in life, due to more limited experiences with the extent of the disparities in levels of living in society as a whole; a poor person's idea of what it means to be "rich" may then be very different to that of a middle or upper income person with a very different frame of reference. This may be correlated with certain attributes of the household.

Recognizing the existence of latent heterogeneity in subjective welfare can also alter the interpretation of some of the regressions found in the literature. Consider, for example, the interesting argument by Knight and Gunatilaka (2012) about the existence of an "hedonic

treadmill” in China. Recall that this was based on a regression of SWL against both answers to the MIQ (taken to proxy “aspiration income”) and actual income (with controls for some other observed covariates, which can be ignored for expository purposes). However, the answers to the MIQ may simply be picking up latent heterogeneity in subjective welfare. Indeed, it can be readily shown that the economic model of the SSPL outlined in section 3 implies that any latent individual characteristic that increases welfare will reduce y_i^{\min} at given own income.³² The MIQ variable may not be measuring aspirations at all but simply acting as a (negative) control for omitted effects on subjective welfare. This will increase precision and (probably) reduce bias in estimating the income effect, which is strong in the Knight and Gunatilaka regressions. But this is clearly a very different interpretation to the idea of an hedonic treadmill.

We have seen how much latent heterogeneity can confound the interpretation of subjective welfare regressions. What can be done about the problem? One possible solution is to control for these differences directly. Psychologists have developed various tests that can be implemented in surveys.³³ However, it is not as yet common to have such data collected in the surveys needed to study the determinants of subjective welfare. The next sub-section considers an econometric method of neutralizing latent heterogeneity.

Addressing latent heterogeneity using panel data

Panel data on subjective welfare and covariates have been proposed as one solution to the problem of latent heterogeneity. If we have longitudinal data on self-assessed welfare (SW) and covariates (X) we postulate a model of the form:

$$SW_{it} = \alpha + \beta X_{it} + \eta_i + \varepsilon_{it} \quad (i=1, \dots, N; t=1, \dots, T) \quad (10)$$

where the error term has two components, a time-invariant effect capturing (say) personality traits, η_i , for which $Cov(\eta_i, X_{it}) \neq 0$, and an innovation error term, ε_{it} with $Cov(\varepsilon_{it}, X_{it}) = 0$.

Thus the non-ignorable heterogeneity is assumed to be entirely captured by η_i . One can question

³² To see this, consider the simplified version of equations (5) and (6) where utility depends on own-income but there is a latent heterogeneity term, so utility is $v(y_i) + \eta_i$. The MIQ solves: $v(y_i^{\min}) + \eta_i = \phi[v(y_i) + \eta_i]$. Differentiating w.r.t. η_i holding own-income constant we find that: $\partial y_i^{\min} / \partial \eta_i = [\phi'(\cdot) - 1] / v'(y_i^{\min})$. The conditions for existence of a unique SSPL in utility space imply that $\phi'(\cdot) < 1$ at least locally. Then $\partial y_i^{\min} / \partial \eta_i < 0$.

³³ The Pearlin scale is an example (Pearlin and Schooler, 1978).

whether this econometric specification adequately captures the way heterogeneity impacts on subjective welfare; I will return to this point after considering some examples.

In what appears to have been the first panel data study of subjective welfare, van de Stadt et al., (1985) modeled a money metric of subjective welfare in the Netherlands. They did not allow for latent individual effects correlated with the covariates, but they did allow for dynamics, by including the lagged subjective welfare measure as a regressor, which picks up at least some of the heterogeneity. More recently, the papers by Winkelmann and Winkelmann (1998), Ravallion and Lokshin (2001) and Senik (2004) allowed for a time-invariant latent individual effect when modeling subjective welfare data. I will discuss the example of the Ravallion and Lokshin paper in more detail as this provides insights into the implications of allowing for latent heterogeneity for standard regression results on the determinants of subjective welfare.

Using panel data for Russia on responses to the ELQ, Ravallion and Lokshin confirmed some of the results in past studies using single cross-sections, but not others. Household income was found to be a highly significant predictor of self-rated economic welfare. Controlling for household income, individual income was a far weaker predictor, suggesting that there is some degree of income pooling. Health shocks lowered subjective economic welfare, at given values of other variables in their model, including incomes. The demographic effects found in cross-sectional studies (notably of household size, at given income per capita) were not found to be robust. The extent of the scale economy of household size in individual subjective welfare suggested by a number of papers in the literature may well reflect latent personality effects on the demographic characteristics of the respondent's household.

In keeping with past work, Ravallion and Lokshin found evidence of an income-compensated welfare cost of unemployment. They argue that most plausible interpretation is that the current welfare loss arises from some combination of lower expected future incomes and more uncertain incomes, the latter arising from greater exposure to uninsured risk outside the formal employment sector. They found that it would take a large gain in current income to compensate for becoming unemployed.

However, their results point to a number of qualifications on past findings about the welfare effects of unemployment. While becoming unemployed entails a large welfare loss, that loss is not fully restored when an unemployed person gets a job, except via the income gain. This

implies a long-term welfare loss from even transient unemployment at given income. It also suggests that high unemployment benefits do not attract people out of work, but they may well discourage a return to work. Also, the unemployment effect was not found to be robust to restricting the sample to those for whom incomes were almost certainly better measured. This suggests that there could well be a bias in past estimates of the (income-compensated) welfare effect of unemployment, arising from the structure of income measurement errors.

These panel data models impose considerable structure on the way in which latent heterogeneity is deemed to matter to subjective welfare, namely that it takes the form of a individual specific time-invariant and additive component of the error term. I turn next to another source of data that offers a more flexible and robust way of addressing concerns about latent heterogeneity in scales.

Using vignettes to address heterogeneity in scales

The essential idea of vignettes in this context is to include in the survey (or a sub-sample) reasonably detailed descriptions of the lives of certain stylized individuals and ask the respondent the same subjective question about those vignettes that was posed about the respondent herself.

Vignettes to have been used in a number of recent studies of subjective data on (*inter alia*) health status, political efficacy and job satisfaction. King et al. (2004) and King and Wand (2007) designed vignettes to establish common points on the heterogeneous reference scales regarding political efficacy in China and Mexico. Kristensen and Johansson (2008) used vignettes in anchoring subjective scales for job satisfaction. Bago d’Uva et al. (2008) used them for correcting self-assessed health data for reporting bias. Vignettes have also been used in testing the competence of doctors (Das et al., 2008). Kapteyn et al. (2008) used vignettes to compare life satisfaction between respondents in the U.S. and the Netherlands.

Beegle, Himelein and Ravallion (BHR) (2012) used vignettes to provide various tests for bias due to latent heterogeneity in individual scales of subjective welfare. Their tests entailed adding vignettes to a national household survey for Tajikistan in 2007. Respondents placed these vignettes on the same six-step ladder used to report their own subjective economic welfare. BHR performed three sets of tests for bias. In their first test, they found some significant covariates for

vignette responses amongst the regressors commonly used to explain subjective welfare. However, the effects were neither very strong nor consistent across different vignettes.

In their second and third tests, BHR explored the extent of bias due to the aforementioned frame of reference effect by embedding vignette rankings by households in standard regressions for subjective welfare. BHR did this in two ways: they included vignettes among the covariates and, to address endogeneity concerns with this approach, BHR also recalibrated the self-assessments of welfare to accord with the heterogeneity in scales revealed by the vignette responses. The striking finding of these tests is that the bias is negligible with respect to the “income effect” on subjective welfare as well as other covariates. Based on this study, the concerns that past uses of subjective economic welfare data are compromised by systematic differences in the meaning given to the scales used appear to be unwarranted.

A third test by BHR involved estimating an explicit model of the heterogeneity in scales, whereby the thresholds are modeled as functions of covariates, assuming that the structure of this heterogeneity is common between “own welfare” and the vignettes.³⁴ As usual, there is assumed to be a continuous latent variable for the respondents’ self-rated welfare, but now the thresholds are explicit functions of a vector of explanatory variables. The extra information on the vignette responses provides the basis for identification, under the assumption that the thresholds for self-assessed welfare are determined identically as for these thresholds in the vignette responses. This is a natural assumption to make. Without the vignettes, identification would only be possible under questionable assumptions about the nonlinearity of the functional forms involved (as discussed in King et al., 2004). Thus BHR were able to model determinants of the thresholds separately to those of the latent continuous variable for subjective welfare.

This approach revealed similar results on the factors influencing subjective welfare, but also threw new light on the heterogeneity in scales. In particular, BHR found that poorer households tend to have a narrower range in the thresholds used in identifying where they lie and where the vignettes lie on the subjective welfare ladder.

³⁴ Here BHR followed an approach developed by King et al., (2004) who proposed an econometric method (“CHOPIT”) for modeling the same type of data use by BHR in which there is an equation for both the latent response of interest and one for the vignettes, with the two equations estimated jointly.

A frame-of-reference effect on subjective welfare is evident in BHR's findings; people with different socioeconomic backgrounds use systematically different scales in responding to questions on their welfare. However, BHR's results do not suggest that this is an important source of bias in past efforts to model the objective determinants of subjective welfare, and for retrieving social subjective poverty lines. This is encouraging for future applications.

Latent heterogeneity in weights

So far the discussion has focused on the problem of heterogeneity in self-assessed welfare, interpreted as either an additive effect on the underlying continuous welfare level or in the ordinal scales used in the survey. However, there is another concern that has received little or no attention in this context.

Recall that one objective of the exercise is to use subjective welfare data to help determine the weights on various non-income dimensions of welfare—to inform choices in constructing a multidimensional index of poverty. It appears very likely that different people have different weights given that non-market goods are included, and so there is no market mechanism to bring marginal rates of substitution into parity across people.

What then does the standard subjective-welfare regression give us? It appears to be widely thought that the regression coefficients can be interpreted as the average weights for the population.³⁵ However, that interpretation requires some rather strong behavioral assumptions. To see why, note first that the differences in weights across the sample will end up in the error term of the regression model. Then a key issue is whether the heterogeneity in the weighted idiosyncratic welfare determinants is correlated with the dimensions of welfare one is interested in. If it is then we have an example of what (in the context of impact evaluations) Heckman et al. (2006) term “essential heterogeneity.” In the present context, let us continue to suppose that subjective welfare is a linear function of X_i , as in equation (9) with $Cov(X_i, \varepsilon_i) = 0$, but now allow the weights to vary across people, so that $\beta_i = \bar{\beta} + \gamma_i$ is the weight for person i and $\bar{\beta}$ is the population mean weight. Then equation (9) becomes:

$$SW_i = \alpha + \beta_i X_i + \varepsilon_i = \alpha + \bar{\beta} X_i + (\varepsilon_i + \gamma_i X_i) \quad (11)$$

³⁵ See, for example, the discussion in Kingdon and Knight (2006).

The composite error term in parentheses contains the heterogeneity in weights. Now it is far from clear what standard subjective welfare regressions are telling us. Even if $Cov(X_i, \varepsilon_i) = 0$, a linear regression model for subjective welfare will not in general deliver the mean weights in the population in large samples. That requires that the interaction effects between the latent differences in weights and the values taken by the covariates are orthogonal to those covariates, i.e., that $Cov(X_i, \gamma_i X_i) = 0$. However, we can easily imagine that people with high weights on some welfare dimension tend to be better endowed with that dimension ($Cov(X_i, \gamma_i X_i) > 0$); this may reflect parental decisions on schooling, say. Then (ignoring all other problems) the regression coefficient will tend to over-estimate the mean weight. Alternatively, suppose that people with low innate attributes (such as health) tend to put a higher value on those attributes; then one could find that $Cov(X_i, \gamma_i X_i) < 0$. We will underestimate the mean weight.

Thus the likely heterogeneity in the weights attached to non-market goods clouds inferences from subjective-welfare regressions. Indeed, while one would presumably like to find weights appropriate to the trade-offs of poor people, it is far from clear that the weights implied by standard regressions are appropriate. One could try to model the heterogeneity in weights by adding interaction effects. This will presumably help, although latent heterogeneity in weights can never be ruled out.

A possible direction for research is to use surveys to try to focus directly on the subjective tradeoffs adopted by respondents. This would require more complicated and difficult survey questions, but a not dissimilar approach has been used with some success for inferring attributes of preferences, such as risk aversion or valuations of environmental factors.

6. Conclusions

Subjective data offer to expand the information set traditionally used for assessing welfare and measuring poverty. Self-assessed welfare is not equated with welfare. Rather, the promise of subjective data is to help identify the weights on the dimensions of welfare for which prices are missing (or unreliable) and in determining the social subjective poverty line—below which people tend to think they are poor, but above which they do not.

Some variables with explanatory power for subjective welfare may not be relevant for measuring poverty. Personality traits, for example, are expected to influence self-assessed welfare but are unlikely to carry much weight in deciding if one person is poorer than another. And even when personality traits are not data, they are likely to influence the observed covariates of self-assessed welfare. In the end, one cannot escape the need for normative judgments about what should be included in a measure of poverty.

The existence of confounding “non-welfare” factors in survey-based measures of subjective welfare is a continuing concern. Practitioners should not presume that these variables are statistically harmless for the purpose at hand. There are a number of reasons for expecting biases in the subjective welfare regressions found in the literature, especially those using cross-sectional data. However, some encouragement can also be found in the recent literature. For example, while there are frame-of-reference effects—whereby the scales in subjective questions are interpreted differently depending on the respondent’s knowledge and experience—the biases turned out to be quite small in the one study to date using vignettes to eliminate the heterogeneity in subjective scales for a developing country.

A potentially important problem that has received no attention (to my knowledge) is latent heterogeneity across individuals in their welfare weights—as must be expected for non-market goods. Those weights may well reflect endowments of the welfare covariates, or the covariates may come to reflect the differing weights. Then we do not know whose trade-offs the subjective-welfare regression is giving us; we would like to use weights favored by poor people, but that cannot be assured. Further research might usefully explore the possibilities of deriving a locally valid estimator using innovative survey or experimental methods to identify personal tradeoffs.

References

- Alkire, Sabina and James Foster, 2007, "Counting and Multidimensional Poverty Measurement," Oxford Poverty and Human Development Initiative, Working Paper 7, University of Oxford.
- Alkire, Sabina and Maria Emma Santos, 2010, "Acute Multidimensional Poverty: A New Index for Developing Countries," Oxford Poverty and Human Development Initiative, Working Paper 38, University of Oxford.
- Angoff, William H., 1993, "Perspectives on Differential Item Functioning Methodology. In Holland, P., Wainer, H. (Eds.) *Differential Item Functioning*. Hillsdale NJ: Lawrence Erlbaum Associates, 3-24.
- Atkinson, Anthony B., 1987, "On the Measurement of Poverty," *Econometrica* 55: 749-64.
- Atkinson, Anthony B. and Francois Bourguignon, 2001, "Poverty and Inclusion from a World Perspective." In Joseph Stiglitz and Pierre-Alain Muet (eds) *Governance, Equity and Global Markets*, Oxford: Oxford University Press.
- Bago d'Uva, T., Van Doorslaer, E., Lindeboom, M., O'Donnell, O., 2008, "Does Reporting Heterogeneity Bias the Measurement of Health Disparities?" *Health Economics* 17(3): 351-375.
- Beegle, Kathleen, Kristen Himelein and Martin Ravallion, 2012, "Frame-of-Reference Bias in Subjective Welfare Regressions." *Journal of Economic Behavior and Organization*, forthcoming.
- Bertrand, Marianne and Sendhil Mullainathan, 2001, "Do People Mean What They Say? Implications for Subjective Survey Data," *American Economic Review, Papers and Proceedings* 91(2): 67-72.
- Bishop, John A., and Feijun Luo, 2006, "Economic Transition and Subjective Poverty in Urban China," *Review of Income and Wealth* 52(4): 625-641.
- Blackorby, Charles and Donaldson, David, 1987, "Welfare Ratios and Distributionally Sensitive Cost-Benefit Analysis," *Journal of Public Economics* 34: 265-90.
- Bourguignon, Francois and Satya Chakravarty, 2003, "The Measurement of Multidimensional Poverty," *Journal of Economic Inequality* 1: 25-49.
- Browning, Martin, 1992, "Children and Household Economic Behavior," *Journal of Economic Literature* 30(3): 1434-75.

- Cantril, Hadley, 1965, *The Pattern of Human Concerns*, New Brunswick: Rutgers University Press.
- Carletto, Gero and Alberto Zezza, 2006, "Being Poor, Feeling Poor: Combining Objective and Subjective Measures of Welfare in Albania," *Journal of Development Studies* 42(5): 739-760.
- Castilla, Carolina, 2010, "Subjective Poverty and Reference-Dependence Income over Time, Aspirations and Reference Groups," UNU-WIDER Working Paper 2010/76.
- Clark, Andrew E., Paul Frijters, and Michael A. Shields, 2008, "Relative Income, Happiness and Utility: An Explanation for the Easterlin Paradox and Other Puzzles," *Journal of Economic Literature* 46(1): 95-144.
- Clark, Andrew E., and Andrew J. Oswald, 1994, "Unhappiness and Unemployment," *Economic Journal* 104: 648-59.
- _____, and _____, 1996, "Satisfaction and Comparison Income," *Journal of Public Economics* 61: 359-381.
- Clark, Andrew E., and Claudia Senik, 2010, "Who Compares to Whom? The Anatomy of Income Comparisons in Europe," *Economic Journal* 120: 573-594.
- Conti, Gabriella and Stephen Pudney, 2011, "Survey Design and the Analysis of Satisfaction," *Review of Economics and Statistics* 93(3): 1087-1093.
- Corazzini, Luca, Lucio Esposito and Francesca Majorano, 2010, "Exploring the Absolutist Vs, Relativist Perception of Poverty Using Cross-Country Questionnaire Survey," *Journal of Economic Psychology*, forthcoming.
- Danziger, S., J. van der Gaag, E. Smolensky, E. and M. Taussig, 1984, "The Direct Measurement of Welfare Levels: How Much does it Take to Make Ends Meet," *Review of Economics and Statistics* 66, 500-5.
- Das, Jishnu, Jeff Hammer, and Ken Leonard, 2008, "The Quality of Medical Advice in Low-Income Countries," *Journal of Economic Perspectives* 22(2): 93-114.
- De Vos, Klaas and Thesia Garner, 1991, "An Evaluation of Subjective Poverty Definitions: Comparing Results from the U.S. and the Netherlands," *Review of Income and Wealth* 37(3): 267-285.
- Easterlin, Richard A., 1974, "Does Economic Growth Improve the Human Lot? Some Empirical

- Evidence,” in P.A. David and W.R. Melvin (eds) *Nations and Households in Economic Growth*, Palo Alto: Stanford University Press.
- Fafchamps, M., and F. Shipli, 2009, “Isolation and Subjective Welfare: Evidence from South Asia.” *Economic Development and Cultural Change*. 57(4), 641-83.
- Ferrer-i-Carbonell, Ada., 2005, “Income and Well-Being: An Empirical Analysis of the Comparison Income Effect.” *Journal of Public Economics*. 89, 997-1019.
- Ferrer-i-Carbonell, Ada and Bernard Van Praag, 2001, “Poverty in Russia,” *Journal of Happiness Studies* 2: 147-172.
- Foster, James, and A. F. Shorrocks, 1988, “Poverty Orderings,” *Econometrica* 56: 173-177.
- Frank, Robert H., 1997, “The Frame of Reference as a Public Good.” *Economic Journal*. 107, 1832-1847.
- Frey, B., Stutzer, A., 2002, “What Can Economists Learn from Happiness Research?” *Journal of Economic Literature*. XL, 402–435.
- Goldberg, L. R., 1993, “The Structure of Phenotypic Personality Traits.” *American Psychologist* 48(1): 26-34.
- Graham, Carol and Stefano Pettinato, 2002, “Frustrated Achievers: Winners, Losers, and Subjective Well-being in New Market Economies,” *Journal of Development Studies* 38(4): 100–140.
- Groedhart, T., V. Halberstadt, A. Kapteyn, A. and B. Van Praag, 1977, “The Poverty Line: Concept and Measurement,” *Journal of Human Resources* 12: 503–20.
- Gustafsson Bjo, Li Shi and Hiroshi Sato, 2004, “Can a Subjective Poverty Line be Applied to China? Assessing Poverty among Urban Residents in 1999,” *Journal of International Development* 16: 1089–1107.
- Hagenaars, Aldi, and Bernard van Praag, 1985, “A Synthesis of Poverty Line Definitions,” *Review of Income and Wealth* 31(2): 139-54.
- Heckman, James, 2011, “Integrating Personality Psychology into Economics,” NBER Working Paper 17378.
- Heckman, James, Serio Urzua and Edward Vytlačil, 2006, “Understanding Instrumental Variables in Models with Essential Heterogeneity,” *Review of Economics and Statistics* 88(3): 389-432.

- Herrera, Javier, Mireille Razafindrakoto and François Roubaud, 2006, "The Determinants of Subjective Poverty: A Comparative Analysis Between Madagascar and Peru," Paris: DIAL Working Paper DT/2006-01.
- Kahneman, D., and Varey, C., 1991, "Notes on the Psychology of Utility," in J. Elster and J. Roemer, eds. *Interpersonal Comparisons of Well-being*. Cambridge: Cambridge University Press.
- Kapteyn, Arie, 1994, "The Measurement of Household Cost Functions: Revealed Preference versus Subjective Measures," *Journal of Population Economics* 7: 333–50.
- Kapteyn, Arie., Kooreman, P. and Willemse, R., 1988, "Some Methodological Issues in the Implementation of Subjective Poverty Definitions," *Journal of Human Resources* 23: 222–42.
- Kapteyn, Arie, Smith, J., Van Soest, A., 2008, "Comparing Life Satisfaction." Working Paper WR-623, Rand Cooperation.
- Kapteyn, Arie and Bernard Van Praag, 1976, "A New Approach to the Construction of Family Equivalence Scales," *European Economic Review* 7: 313-335.
- Kilpatrick, Robert, 1973, "The Income Elasticity of the Poverty Line." *Review of Economics and Statistics* 55(3): 327-32.
- King, G., Murray, C., Salomon, J., and Tandon, A., 2004, "Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research," *American Political Science Review*. 98(1): 191-207.
- King, G., Wand, J., 2007, "Comparing Incomparable Survey Responses: Evaluating and Selecting Anchoring Vignettes," *Political Analysis*. 15(1): 46-66.
- Kingdon, Geeta and John Knight, 2006, "Subjective Well-being Poverty vs. Income Poverty and Capabilities Poverty?" *Journal of Development Studies* 42(7): 1199-1224.
- _____ and _____, 2007, "Community, Comparisons and Subjective Well-Being in a Divided Society," *Journal of Economic Behavior and Organization* 64(1): 69-90.
- Knight, John and Ramani Gunatilaka, 2010, "The Rural-Urban Divide in China: Income but not Happiness?" *Journal of Development Studies* 46(3): 506-534.
- _____ and _____, 2012, "Income, Aspirations and the Hedonic Treadmill in a Poor Society," *Journal of Economic Behavior and Organization*, forthcoming.

- Kristensen, Nicolai and E. Johansson, 2008, "New Evidence on Cross-Country Differences in Job Satisfaction Using Vignettes," *Labor Economics* 15(1): 96-117.
- Kristensen, Nicolai and Niels Westergaard-Nielsen, 2007, "Reliability of Job Satisfaction Measures," *Journal of Happiness Studies* 8: 273-292.
- Krueger, Alan B., and David Schkade, 2008, "The Reliability of Subjective Well-being Measures," *Journal of Public Economics* 92(8-9): 1833-1845.
- Lanjouw, Peter, and Martin Ravallion, 1995, "Poverty and Household Size," *Economic Journal* 105: 1415-1435.
- Lokshin, Michael and Martin Ravallion, 2008, "Testing for an Economic Gradient in Health Status Using Subjective Data," *Health Economics* 17: 1237-1259.
- Lokshin, Michael, Nithin Umapathi and Stefano Paternostro, 2006, "Robustness of Subjective Welfare Analysis in a Poor Developing Country: Madagascar 2001," *Journal of Development Studies* 42(4): 559-591.
- Luttmer, Erzo F.P., 2005, "Neighbors as Negatives: Relative Earnings and Well-Being," *Quarterly Journal of Economics* 120(3): 963-1002.
- Morris, M. D., 1979, *Measuring the Condition of the World's Poor: The Physical Quality of Life Index*, Washington, D.C.: Overseas Development Council.
- Mueller, Gerrit and Erik Plug, 2006, "Estimating the Effect of Personality on Male and Female Earnings," *Industrial and Labor Relations Review* 60(1): 3-22.
- Oswald, Andrew, 1997. "Happiness and Economic Performance," *Economic Journal* 107: 1815-31.
- Pearlin, Leonard and Carmi Schooler, 1978, "The Structure of Coping," *Journal of Health and Social Behaviour* 19: 2-21.
- Pollak, Robert, 1991, "Welfare Comparisons and Situation Comparisons," *Journal of Econometrics* 50: 31-48.
- Pollak, Robert and Wales, Terence, 1979, "Welfare Comparison and Equivalence Scale," *American Economic Review* 69: 216-21.
- Pradhan, Menno and Martin Ravallion, 2000, "Measuring Poverty using Qualitative Perceptions of Consumption Adequacy," *Review of Economics and Statistics* 82: 462-71.
- Ravallion, Martin, 1990, "Income Effects on Undernutrition," *Economic Development and Cultural Change*, 38: 489-516.

- _____, 1994, *Poverty Comparisons*. Harwood Academic Press, Chur: Switzerland.
- _____, 1998, *Poverty Lines in Theory and Practice*. Living Standards Measurement Study Paper 133. Washington DC: World Bank.
- _____, 2003, "The Debate on Globalization, Poverty and Inequality: Why Measurement Matters," *International Affairs*, 79(4): 739-754.
- _____, 2008b, "Poverty Lines," in *The New Palgrave Dictionary of Economics*, 2nd Edition, Larry Blume and Steven Durlauf (eds) London: Palgrave Macmillan.
- _____, 2008b, "On the Welfarist Rationale for Relative Poverty Lines," in Kaushik Basu and Ravi Kanbur (eds) *Social Welfare, Moral Philosophy and Development: Essays in Honour of Amartya Sen's Seventy Fifth Birthday*, Oxford: Oxford University Press.
- _____, 2010, "The Developing World's Bulging (but Vulnerable) Middle Class," *World Development* 38(4): 445-454.
- _____, 2011, "On Multidimensional Indices of Poverty," *Journal of Economic Inequality*, 9(2): 235-248.
- _____, 2012a, "Troubling Tradeoffs in the Human Development Index," *Journal of Development Economics*, forthcoming.
- _____, 2012b, "Poverty Lines Across the World," Chapter 3 of the *Oxford Handbook of the Economics of Poverty*, edited by Philip N. Jefferson, Oxford University Press.
- Ravallion, Martin and Shaohua Chen, 2011, "Weakly Relative Poverty," *Review of Economics and Statistics* 93(4): 1251-1261.
- Ravallion, Martin and Michael Lokshin, 2000, "Who Wants to Redistribute? The Tunnel Effect in 1990s Russia," *Journal of Public Economics* 76(1): 87-104.
- _____ and _____, 2001, "Identifying Welfare Effects from Subjective Questions," *Economica*, 68: 335-357.
- _____ and _____, 2002, "Self-assessed Economic Welfare in Russia," *European Economic Review* 46: 1453-1473.
- _____ and _____, 2006, "Testing Poverty Lines", *Review of Income and Wealth* 52(3): 399-421.
- _____ and _____, 2010, "Who Cares about Relative Deprivation?" *Journal of Economic Behavior and Organization*, 73(2): 171-185.

- Rojas, Mariano, 2007, A Subjective Well-being Equivalence Scale for Mexico: Estimation and Poverty and Income-distribution Implications, *Oxford Development Studies* 35(3): 273-293.
- Scitovsky, Tibor, 1978, *The Joyless Economy*. Oxford: Oxford University Press.
- Sen, Amartya, 1983, "Poor, Relatively Speaking," *Oxford Economic Papers* 35(2):153-69.
- _____, 1985, *Commodities and Capabilities*. Amsterdam: North-Holland.
- _____, 1992, *Inequality Rexamined*, Oxford: Oxford University Press.
- Senik, Claudia, 2004, When Information Dominates Comparison: Learning from Russian Subjective Panel Data. *Journal of Public Economics*. 88, 2099-2123.
- Slesnick, Daniel T., 1993, "Gaining Ground: Poverty in the Postwar United States," *Journal of Political Economy* 101(1): 1-38.
- Solnick, Sara and David Hemenway, 1998, "Is More Always Better? A Survey on Positional Concerns," *Journal of Economic Behavior and Organization* 37: 373-83.
- Stevenson, Betsey and Justin Wolfers, 2008, "Economic Growth and Subjective Well-being: Reassessing the Easterlin Paradox," NBER Working Paper 14282, National Bureau of Economic Research, Cambridge, Massachusetts.
- Taddesse, Mekonnen and Abebe Shimeles, 2005, "Perceptions of Welfare and Poverty: Analysis of the Qualitative Responses of Urban Households," in Arne Bigsten, Abebe Shimeles and Bereket Kebede (eds), *Poverty, Income Distribution and Labour Markets in Ethiopia*, Uppsala, Sweden: Nordic Africa Institute.
- Taylor, Mark P., 2006, "Tell me why I don't like Mondays: Investigating Day of the Week Effects on Job Satisfaction and Psychological Well-being," *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 169(1): 127-142.
- Theodossiou, I., 1998, "The Effects of Low-Pay and Unemployment on Psychological Well-Being: A Logistic Regression Approach, *Journal of Health Economics* 17: 85-104.
- Townsend, Peter, 1979, *Poverty in the United Kingdom*. Harmondsworth, Middlesex: Penguin Books.
- Tsui, Kai-Yuen, 2002, "Multidimensional Poverty Indices," *Social Choice and Welfare* 19: 69-93.
- Tummers, Martijn, 1994, "The Effect of Systematic Misperception of Income on the Subjective

- Poverty Line,” in Richard Blundell, Ian Preston and Ian Walker (eds) *The Measurement of Household Welfare*, Cambridge, England: Cambridge University Press.
- United Nations Development Program (UNDP), 2010, *Human Development Report*, New York: UNDP.
- Van Praag, Bernard, 1968, *Individual Welfare Functions and Consumer Behavior*. Amsterdam: North-Holland.
- _____, 1994, “Ordinal and Cardinal Utility: An Integration of the Two Dimensions of the Welfare Concept,” in Richard Blundell, Ian Preston and Ian Walker (eds) *The Measurement of Household Welfare*, Cambridge, England: Cambridge University Press.
- van de Stadt, Huib, Arie Kapteyn, and Sara van de Geer, 1985, “The Relativity of Utility: Evidence from Panel Data,” *Review of Economics and Statistics* 67: 179-187.
- Wagle, Udaya, 2008, *Multidimensional Poverty Measurement: Concepts and Applications*, New York: Springer.
- Winkelmann, Liliana and Rainer Winkelmann, 1998, “Why Are the Unemployed So Unhappy? Evidence from Panel Data,” *Economica* 65: 1-15.
- Zheng, Buhong, 1997, “Aggregate Poverty Measures,” *Journal of Economic Surveys* 11(2): 123-140.